A Stability Condition for Certain Bilinear Systems

Junghsi Lee and V. J. Mathews

Abstract—**This correspondence derives a simple sufficient condition for the output of a discrete-time, time-invariant bilinear system to be bounded whenever the input signal to the system is bounded by a finite constant.**

I. Introduction

The correspondence considers real-valued, discrete-time and time-invariant bilinear systems whose output $y(n)$ is related to the input $x(n)$ through the nonlinear difference equation

$$y(n) = \sum_{i=0}^{N_1} a_i x(n-i) + \sum_{i=1}^{N_2} b_i y(n-i) + \sum_{i=0}^{N_3}\sum_{j=1}^{N_4} c_{i,j} x(n-i) y(n-j). \quad (1)$$

In (1), a_i's, b_i's, and $c_{i,j}'$s represent the coefficients of the bilinear system model. Such models are very attractive in many practical applications since a large class of nonlinear systems can be adequately and efficiently approximated with finite number of coefficients [2], [7]. Because of this, bilinear system models have found applications in several areas including communications, control systems, population models, etc. Several applications and properties of bilinear system models are reviewed in [3], [7]. Adaptive nonlinear filters using bilinear system models have also been developed recently [1], [5], [6].

In spite of their usefulness, there is one major problem with bilinear system models in that most bilinear systems are inherently unstable. By this, we mean that it is possible to find bounded input signals that can drive the output signal to become unbounded for almost any given bilinear system. However, there is a class of input signals associated with most bilinear systems such that they generate well behaved and useful output signals whenever their inputs belong to this class.

The purpose of this correspondence is to develop a sufficient condition that guarantees that the output of a bilinear system is bounded whenever its input signal is bounded by a finite constant (say, M_x). This condition will enable us to define a subset of the useful input signals associated with a given bilinear system, and consequently be more confident about the usefulness of the system model in the application involved.

The problem of deriving the "input-dependent" stability condition as defined above is closely related to the problem of deriving stationarity conditions for a bilinear system series. However, this problem is also very difficult, and simple solutions are available only for the simplest bilinear time series models [8]. The conditions for the most general bilinear time series models similar to that in (1) require finding the eigenvalues of a $r^2 s \times r^2 s$ matrix where $r = \max\{N_2, N_4\}$ and $s = \max\{N_1, N_3, N_4\}$. Therefore, they are computationally extremely complex. In contrast, our approach is applicable for the most general discrete-time, time-invariant bilinear systems and requires one to find the zeros of an N_2th order polynomial.

The remainder of this correspondence is organized as follows. The next section derives a simple sufficient condition for the output of (1) to be bounded whenever the input signal to the system is bounded by a finite constant. The concluding remarks are made in Section III.

II. The Stability Condition

For simplicity of notation, some time-invariant operators are defined as follows

$$A(q) = \sum_{i=0}^{N_1} a_i q^{-i} \quad (2)$$

$$B(q) = \sum_{i=1}^{N_2} b_i q^{-i} \quad (3)$$

$$C(q) = \sum_{i=0}^{N_3}\sum_{j=1}^{N_4} c_{i,j}[q^{-i}, q^{-j}] \quad (4)$$

where $q^{-i}x(n) = x(n-i)$ and $[q^{-i}, q^{-j}][x(n), y(n)] = x(n-i)y(n-j)$. Using these operations, the bilinear model in (1) can be represented as

$$(1 - B(q))y(n) = A(q)x(n) + C(q)[x(n), y(n)]. \quad (5)$$

The main result of this correspondence is contained in the following theorem:

Theorem: For the bilinear system model represented by (5), let $q_1, q_2, \cdots,$ and q_{N_2} denote the zeros of the polynomial $q^{N_2}(1 - B(q))$. Given a real, positive number M_x

$$\begin{cases} |q_i| < 1, \text{ for } i = 1, 2, \cdots, N_2, \text{and} \\ M_x \sum_{i=0}^{N_3}\sum_{j=1}^{N_4} |c_{i,j}| < \prod_{i=1}^{N_2}(1 - |q_i|) \end{cases} \quad (6)$$

constitute a sufficient condition for every $x(n)$ bounded by M_x to produce a bounded output $y(n)$. Furthermore, $y(n)$ is bounded by

$$|y(n)| \le \frac{\beta M_x \sum_{i=0}^{N_1} |a_i|}{1 - \left\{\beta M_x \sum_{i=0}^{N_3}\sum_{j=1}^{N_4} |c_{i,j}|\right\}} \quad (7)$$

where

$$\beta = \frac{1}{\prod_{i=1}^{N_2}(1 - |q_i|)}. \quad (8)$$

It should be noted that in the theorem $q^{N_2}(1 - B(q))$ represents an N_2th order polynomial rather than an operator. As long as it does not cause any confusion, this notation will be used throughout the correspondence. Also note that when there is no nonlinearity in the system, (5) reduces to a linear recursive system, and condition (6) reduces to the well-known necessary and sufficient condition for the resulting linear system to be stable in the bounded-input bounded-output sense.

In the proof given below, we will first define a sequence of system with outputs $y_k(n), k = 1, 2, \cdots, n+1$, then show that $y(n) = \sum_{k=1}^{n+1} y_k(n)$ is the unique solution to the system (5), and finally prove that (6) constitute a sufficient condition for $y(n)$ to be bounded.

Manuscript received September 2, 1992; revised August 2, 1993. This work was supported in part by the National Science Foundation under Grant MIP-8922146, by a Department Grant from the IBM Corporation, and by Project No. 37H3100 conducted by the ITRI under sponsorship of the Minister of Economic Affairs, R.O.C. The associate editor coordinating the review of this paper and approving it for publication was Prof. José A. R. Fonollosa.

J. Lee is with the Advanced Technology Center, CCL, ITRI, Taiwan, R.O.C.

V. J. Mathews is with the Department of Electrical Engineering, University of Utah, Salt Lake City, UT 84112 USA.

IEEE Log Number 9401269.

Proof: For simplicity, we assume that system (5) is initially at rest. Define a sequence of systems as follows

$$\begin{cases} (1-B(q))y_1(n) &= A(q)x(n), \\ (1-B(q))y_k(n) &= C(q)[x(n), y_{k-1}(n)] \\ \quad \text{for } k &= 2,3,\cdots,n+1 \\ y_1(T)=y_2(T) &= \cdots = y_{n+1}(T) = 0 \text{ for } T \le -1. \end{cases} \tag{9}$$

It is straightforward to show that

$$y_1(n) = \left\{ \prod_{m=1}^{N_2} \left\{ \sum_{l=0}^{n} (q_m)^l q^{-l} \right\} \right\} A(q)x(n) \tag{10}$$

and

$$y_k(n) = \left\{ \prod_{m=1}^{N_2} \left\{ \sum_{l=0}^{n} (q_m)^l q^{-l} \right\} \right\} C(q)[x(n), y_{k-1}(n)] \tag{11}$$

for $k = 2,3,\cdots,n+1$. In Appendix we show that $y(n) = \sum_{k=1}^{n+1} y_k(n)$ is the unique solution to the system (5).

Now for $x(n)$ bounded by M_x, assume that all the zeros of $q^{N_2}(1-B(q))$ are inside the unit circle and $\beta M_x \sum_{i=0}^{N_3}\sum_{j=1}^{N_4} |c_{i,j}| < 1$. It follows from (10) that

$$|y_1(n)| \le \beta M_x \sum_{i=0}^{N_1} |a_i|. \tag{12}$$

We can then write

$$|C(q)[x(n), y_1(n)]| \le \left(M_x \sum_{i=0}^{N_3}\sum_{j=1}^{N_4} |c_{i,j}| \right)\left(\beta M_x \sum_{i=0}^{N_1} |a_i| \right). \tag{13}$$

Now, it is easy to conduct that $y_2(n)$ is bounded by

$$|y_2(n)| \le \left(\beta M_x \sum_{i=0}^{N_3}\sum_{j=1}^{N_4} |c_{i,j}| \right)\left(\beta M_x \sum_{i=0}^{N_1} |a_i| \right). \tag{14}$$

In fact, it can be shown that

$$|y_k(n)| \le \left(\beta M_x \sum_{i=0}^{N_3}\sum_{j=1}^{N_4} |c_{i,j}| \right)^{k-1}\left(\beta M_x \sum_{i=0}^{N_1} |a_i| \right) \tag{15}$$

for $k = 3,4,\ldots,n+1$.

It is straightforward to see that

$$\begin{aligned} |y(n)| &\le \sum_{k=1}^{n+1} |y_k(n)| \\ &\le \beta M_x \sum_{i=0}^{N_1} |a_i| \frac{1 - \left\{ \beta M_x \sum_{i=0}^{N_3}\sum_{j=1}^{N_4} |c_{i,j}| \right\}^{n+1}}{1 - \left\{ \beta M_x \sum_{i=0}^{N_3}\sum_{j=1}^{N_4} |c_{i,j}| \right\}} \\ &\le \frac{\beta M_x \sum_{i=0}^{N_1} |a_i|}{1 - \left\{ \beta M_x \sum_{i=0}^{N_3}\sum_{j=1}^{N_4} |c_{i,j}| \right\}}. \end{aligned} \tag{16}$$

This completes the proof. *Q.E.D.*

Remark 1: In fact, $y(n) = \sum_{k=1}^{n+1} y_k(n)$ is the Volterra series expansion of $y(n)$ in terms of $x(n)$, and $y_k(n)$ is the kth order Volterra term.

Remark 2: The main result of this paper provides a method for guaranteeing that a bilinear system model is "well-behaved". Suppose that the zeros of $q^{N_2}(1-B(q))$ are all bounded by 1. By limiting the amplitude of the input signals to be bounded by

$$\left(\prod_{i=1}^{N_2} (1 - |q_i|) \right) \Bigg/ \left(\sum_{i=0}^{N_3}\sum_{j=1}^{N_4} |c_{i,j}| \right)$$

we can force the output to be bounded also.

III. Concluding Remarks

The paper presented a sufficient condition for a discrete-time bilinear system to produce bounded outputs whenever the input signal is bounded by some finite constant. The most attractive aspect of the condition is that its complexity is only proportional to that of evaluating the zeros of an N_2th-order polynomial. Previously available techniques were applicable only for extremely simple systems or they required calculation of the eigenvalues of very large matrices. The simplicity of our approach is obvious from the above statements.

Although the idea of obtaining a Volterra series expansion from a bilinear difference equation is not new (for example, see [4]), and some researchers have considered the stability of nonlinear systems by studying the convergence of the Volterra series also, the results we have are new. Such stability conditions are essential for making the bilinear system models applicable in a large class of applications. However, our result represents only a sufficient condition and further work needs to be done for developing necessary and sufficient conditions.

Even though the condition is developed for time-invariant systems, one could, in practice, apply it to monitor the stability of time-varying systems also. An on-line adaptive version of this stability condition that makes use of an adaptive polynomial factorization algorithm [9] is discussed in [6]. Reference [6] also discusses several other stability conditions for more restricted classes of bilinear systems.

Appendix

In this appendix we will show that $y(n) = \sum_{k=1}^{n+1} y_k(n)$ is the unique solution to the system (5). We first show that $y_k(L) = 0$ for $L \le k-2$ using induction. Because $y_1(L) = 0$ for $L \le -1$, this statement is true for $k = 1$. Assume that it holds for k greater than 1 and less than $n+1$. Therefore, we have $y_n(L) = 0$ for $L \le n-2$. Now consider $k = n+1$. By using (11) and the fact that $y_n(L-j) = 0$ for $L-j \le n-2$, it follows that $y_{n+1}(L) = 0$ for $L-1 \le n-2$, i.e., $L \le (n+1)-2$.

Now that $y_{n+1}(n-j) = 0$ for $j \ge 1$, it follows immediately that

$$\begin{aligned} C(q)[x(n), y_{n+1}(n)] &= \sum_{i=0}^{N_3}\sum_{j=1}^{N_4} c_{i,j} x(n-i) y_{n+1}(n-j) \\ &= 0. \end{aligned} \tag{A.1}$$

Using (9) and (A.1), we can write

$$\begin{aligned} (1-B(q)) \sum_{k=1}^{n+1} y_k(n) &= A(q)x(n) + C(q)\left[x(n), \sum_{k=1}^{n} y_k(n) \right] \\ &= A(q)x(n) + C(q)\left[x(n), \sum_{k=1}^{n+1} y_k(n) \right]. \end{aligned} \tag{A.2}$$

Therefore, $y(n) = \sum_{k=1}^{n+1} y_k(n)$ is a solution to (5).

Now assume that $\dot{y}(n)$ and $\ddot{y}(n)$ are two solutions to (5). Then

$$(1-B(q))(\dot{y}(n) - \ddot{y}(n)) = C(q)[x(n), \dot{y}(n) - \ddot{y}(n)]. \tag{A.3}$$

Since by assumption, $\dot{y}(n) = \ddot{y}(n) = 0$ for $n < 0$, it follows by induction that $\dot{y}(n) = \ddot{y}(n)$ for $n \ge 0$ also. Because $y(n) = \sum_{k=1}^{n+1} y_k(n)$ is a solution to (5), it also has to be the unique solution. *Q.E.D.*

References

[1] H. K. Baik and V. J. Mathews, "Adaptive lattice bilinear filters," *IEEE Trans. Signal Processing*, vol. 41, no. 6, pp. 2033–2047, June 1993.

[2] R. W. Brockett, "Volterra series and geometric control theory," *Automatica*, vol. 12, pp. 167–176, 1976.
[3] C. Bruni, G. Di Pillo, and G. Koch, "Bilinear systems: An appealing class of 'nearly linear systems' in theory and applications," *IEEE Trans. Automat. Contr.*, vol. AC-19, no. 4, pp. 334–348, Aug. 1974.
[4] H. Diaz and A. A. Desrochers, "Modeling of nonlinear discrete-time systems from input-output data," *Automatica*, vol. 24, no. 5, pp. 629–641, 1988.
[5] F. Fnaiech and L. Ljung, "Recursive identification of bilinear systems," *Int. J. Contr.*, vol. 45, no. 2, pp. 453–470, 1987.
[6] J. Lee, "Adaptive polynomial filtering algorithms," Ph.D. dissertation, University of Utah, Salt Lake City, 1992.
[7] R. R. Mohler and W. J. Kolodziej, "An overview of bilinear system theory and applications," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-10, no. 10, pp. 683–688, Oct. 1980.
[8] T. Pham-Dinh and L. T. Tran, "On the first-order bilinear time series model," *J. Appl. Probability*, vol. 18, pp. 617–627, 1981.
[9] D. Starer and A. Nehorai, "Adaptive polynomial factorization by coefficient matching," *IEEE Trans. Signal Processing*, vol. 39, no. 2, pp. 527–530, Feb. 1991.

Application of Alternating Convex Projection Methods for Computation of Positive Toeplitz Matrices

Karolos M. Grigoriadis, Arthur E. Frazho, and Robert E. Skelton

Abstract—**In this correspondence, we use alternating convex projection techniques to compute the closest positive definite Toeplitz matrix that satisfies certain inequality constraints to a specified symmetric matrix. Some applications to signal processing and control problems are discussed.**

I. Introduction

Several problems in digital signal processing and control theory require the computation of a positive definite Toeplitz matrix that closely approximates a given symmetric matrix. For example, in the power spectral estimation of a wide-sense stationary process from a finite number of data, the matrix T_o formed from the estimated autocorrelation coefficients is often not a positive definite Toeplitz matrix [1]. In this case, we are interested to obtain the nearest positive definite Toeplitz matrix to the estimated matrix T_o. In control theory, the Gramian assignment problem for discrete-time single input systems requires the computation of a positive definite Toeplitz matrix, which also satisfies certain inequality constraints [2].

In this correspondence, we use alternating convex projection techniques to solve the following two types of problems: first, to find the positive definite Toeplitz matrix nearest to a given symmetric matrix (*optimization problem*) and second, to find a positive definite Toeplitz matrix that satisfies certain inequality constraints (*feasibility problem*). Alternating projection techniques have been used successfully in image restoration problems [3]. Their use in control design problems was introduced in [4].

Manuscript received January 28, 1993; revised September 14, 1993. This work was supported by NSF Grant 9201393-ECS and ARO Grant DAAL-03-92-G-0027. The associate editor coordinating the review of this paper and approving it for publication was Prof. Georgios B. Giannakis.

The authors are with the School of Aeronautics and Astronautics, Purdue University, W. Lafayette, IN 47907 USA.

IEEE Log Number 9401279.

II. Preliminaries

To begin, let $\mathcal{H}$ be the real Hilbert space of all $n \times n$ symmetric matrices whose inner product is defined by $(T, S) = \text{trace}(TS)$, for T and S in $\mathcal{H}$. In this correspondence, we consider the following optimization problem:

$$\inf\|T_o - T\| \text{ subject to } T \in \mathcal{C} = \mathcal{C}_1 \cap \mathcal{C}_2 \cap \cdots \cap \mathcal{C}_m \qquad (2.1)$$

where T_o is a specified matrix in $\mathcal{H}$, and $\mathcal{C}_i$ are closed convex subsets of $\mathcal{H}$. For example, $\mathcal{C}$ can be the set of all positive definite Toeplitz matrices. In that case, the optimization problem (2.1) seeks the closest positive definite Toeplitz matrix to a given matrix T_o. However, we will treat the problem in a more general setting, and we will also include some other constraints. Note that the optimization problem (2.1) was examined in [4] as well in the context of covariance control design, where the convex sets $\mathcal{C}_i$ correspond to closed-loop covariance matrix assignability constraints and multiple inequality constraints on output variances.

Throughout this correspondence, P_i is the orthogonal projection from $\mathcal{H}$ onto the convex set $\mathcal{C}_i$. Recall that P_i is the possibly nonlinear operator defined by $P_i x = \hat{x}$, where x is in $\mathcal{H}$, and $\hat{x}$ is the unique element in $\mathcal{C}_i$ solving the optimization problem $\|x - \hat{x}\| = \inf\{\|x - y\| : y \in \mathcal{C}_i\}$. Moreover, by the projection theorem, $\hat{x} = P_i x$ is the unique element in $\mathcal{C}_i$ satisfying the following "minimum principle" condition:

$$(x - \hat{x}, y - \hat{x}) \leq 0 \quad (\text{for all } y \in \mathcal{C}_i). \qquad (2.2)$$

Now, let us introduce the convex sets $\mathcal{C}_i$ in which we are interested. To this end, recall that a Toeplitz matrix T is an $n \times n$ matrix whose entries $\{t_{ij}\}$ satisfy $t_{ij} = t_{j-i}$ for all i and j. Now, let $\mathcal{C}_1$ be the n-dimensional subspace of $\mathcal{H}$ defined by the symmetric Toeplitz matrices, that is

$$\mathcal{C}_1 = \{T \in \mathcal{H} : T \text{ is Toeplitz}\}. \qquad (2.3)$$

It is well known that the orthogonal projection P_1 onto $\mathcal{C}_1$ is the linear operator on $\mathcal{H}$ given by

$$P_1 A = \begin{bmatrix} t_0 & t_1 & \dots & t_{n-1} \\ t_1 & t_0 & \dots & t_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n-1} & t_{n-2} & \dots & t_0 \end{bmatrix} \quad (A \in \mathcal{H}) \qquad (2.4)$$

where $A = \{a_{ij}\}$ and the entries $\{t_0, t_1, \cdots, t_{n-1}\}$ of $P_1 A$ are computed by

$$t_0 = \frac{1}{n}\sum_{i=1}^{n} a_{ii}, \quad t_1 = \frac{1}{n-1}\sum_{i=1}^{n-1} a_{ii+1},$$
$$t_2 = \frac{1}{n-2}\sum_{i=1}^{n-2} a_{ii+2}, \ldots, t_{n-1} = a_{1n}. \qquad (2.5)$$

For the optimization problem (2.1), the second convex constrained set in which we are interested is

$$\mathcal{C}_2 = \{T \in \mathcal{H} : T \geq Q\} \qquad (2.6)$$

where Q is a specified positive definite matrix in $\mathcal{H}$. It is well known [5] that the orthogonal projection P_2 from $\mathcal{H}$ onto $\mathcal{C}_2$ is the nonlinear operator given by

$$P_2 A = U\Lambda_+ U^* + Q \qquad (A \in \mathcal{H}) \qquad (2.7)$$

where $*$ denotes transpose. Here, $A - Q = U\Lambda U^*$ is the eigenvalue-eigenvector decomposition of $A - Q$, where U is an $n \times n$ orthogonal